Modelling Paired Comparisons with

The prefmod Package

Regina Dittrich & Reinhold Hatzinger
Institute for Statistics and Mathematics, WU Vienna

Paired Comparisons

- method of data collection
- given a set of J items

- individuals are asked to judge pairs of objects

j preferred to k

k preferred to j

- aim is to rank objects into a preference order
- obtain an overall ranking of the objects

Overview

- LLBT models: loglinear Bradly-Terry models

- Basic LLBT

- Extended LLBT
  undecided response
  subject covariates
  object specific covariates

- Pattern Models
  Paired comparison −▷ pattern models
  Ranking −▷ pattern models
  Rating −▷ pattern models

The Basic Bradley-Terry Model (BT)

for the each comparison (jk) of object j to object k we observe:

- $n_{(j \succ k)}$ . . . the number of times j is preferred to k
- $n_{(k \succ j)}$ . . . the number of times k is preferred to j

$$\boxed{N_{(jk)} = n_{(j \succ k)} + n_{(k \succ j)}}$$

total number of responses to comparison (jk)

the probability that j is preferred to k in comparison (jk)

$$\boxed{P(j \succ k) = \frac{\pi_j}{\pi_j + \pi_k}}$$

π's are a called *worth parameters* and are non-negative numbers describing the location of the objects

The Basic Loglinear BT Model (LLBT)

the model can be formulated as a log-linear model following the usual Multinomial / Poisson - equivalence.

the expected value $m_{(j \succ k)}$ of $n_{(j \succ k)}$ is $\boxed{m_{(j \succ k)} = N_{(jk)} p_{(j \succ k)}}$

$$\boxed{P(j \succ k) = \frac{\pi_j}{\pi_j + \pi_k} = c_{(jk)} \frac{\sqrt{\pi_j}}{\sqrt{\pi_k}}}$$

where $c_{(jk)}$ is constant for a given comparison

then our basic paired comparison model for one comparison is

$$\boxed{\ln m_{(j \succ k)} = \mu_{(jk)} + \lambda_j - \lambda_k}$$

λ's are the object parameters
μ's are nuisance parameters

this model formulation is feasible for further extensions

LLBT with Undecided Response

Using the respecification of the probabilities suggested by Davidson and Beaver (1977):

the LLBT model formulas for the comparison (jk) are now:

$$\ln m_{(j \succ k)} = \mu_{(jk)} + \lambda_j - \lambda_k$$
$$\ln m_{(k \succ j)} = \mu_{(jk)} - \lambda_j + \lambda_k$$
$$\ln m_{(j = k)} = \mu_{(jk)} \qquad\qquad + \gamma$$

where γ is the parameter for undecided response (could also be $\gamma_{(jk)}$)

λ's are the object parameters
μ's are nuisance parameters

terms and relations

- relation between π and λ:

$$\boxed{\lambda_j = \ln \sqrt{\pi_j}}$$

$$\boxed{\pi_j = \exp 2\lambda_j}$$

- identifiability of πs is obtained by the restriction $\pi_J = 1$ via $\lambda_J = 0$
- the worth parameters are calculated by

$$\boxed{\pi_j = \frac{\exp(2\lambda_j)}{\sum_j \exp(2\lambda_j)}, \; j = 1, 2, \ldots, J}$$

where $\sum_j \pi_j = 1$

Example: CEMS exchange programme

students of the WU can study abroad visiting one of currently 17 CEMS universities

aim of the study:

- preference orderings of students for different locations
- identify reasons for these preferences

data:

- PC-responses about their choices of 6 selected CEMS universities for the semester abroad (London, Paris, Milan, Barcelona, St.Gall, Stockholm)
- answer: *can not decide* was allowed
- several covariates (e.g., gender, working status, language abilities, etc.)

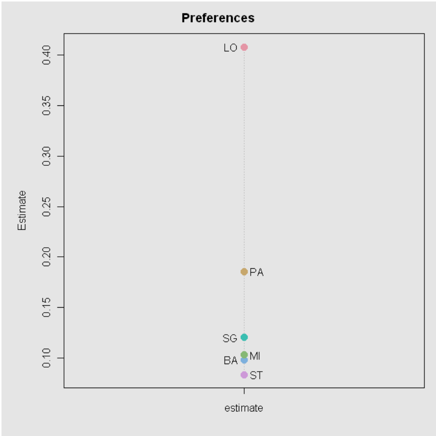
direct estimation: **Function:** `pattPC.fit()`

▷ fit basic LLBT model for PC with undecided

```
> m3 <- llbtPC.fit(cpc, nitems = 6, undec = TRUE, obj.names = cities)
```

Calculate worth and plot using `llbt.worth(), plotworth()`

```
> worth3 <- llbt.worth(m3)
> plotworth(worth3)
```



Subject Covariates

Are the preference orderings different for different groups of subjects?

For one subject covariate on s levels we have now

$$\ln m_{(j \succ k)|s} = \mu_{(jk)s} + \lambda_s^S + (\lambda_j^{O_j} + \lambda_{js}^{O_j S}) - (\lambda_k^{O_k} + \lambda_{ks}^{O_k S})$$

where

λ^O	object parameters (for subject baseline group)
λ^{OS}	interaction parameter between objects and subject category
λ_s^S	fixing the margin for category s of covariate S (nuisance)
μ's	nuisance parameters

Options for: `llbtPC.fit(): formel, elim`

▷ fit model for `SEX*WORK`

```
> msw <- llbtPC.fit(cpc, nitems = 6, undec = TRUE, formel = ~SEX *
+     WORK, elim = ~SEX * WORK, obj.names = cities)
```

Options:

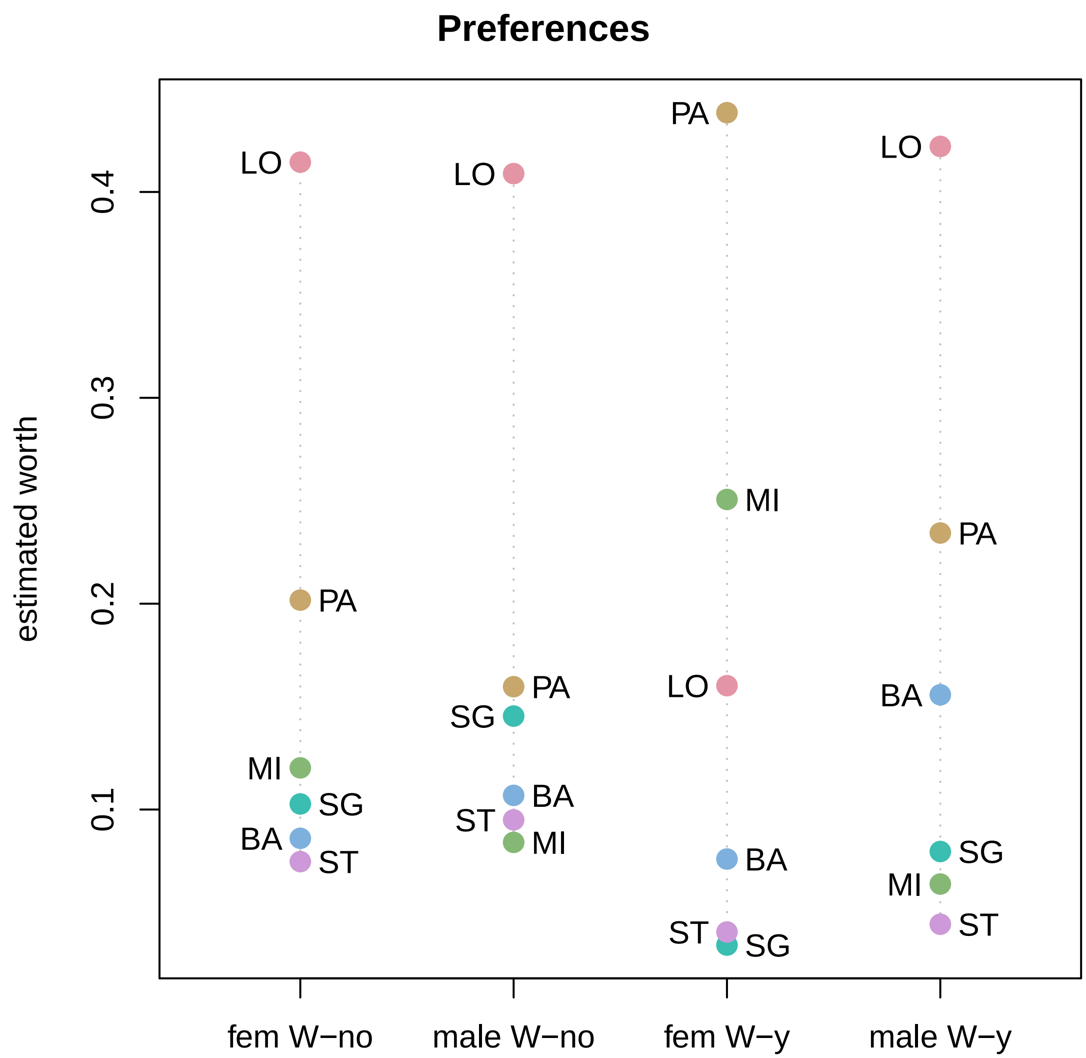
formel $=\sim$ SEX $*$ WORK	model is OBJ $+$ OBJ : (SEX $*$ WORK) OBJ is (LO+PA+M+SG+BA+ST)
elim $=\sim$ SEX $*$ WORK	defines maximal table

▷ now we can easily generate worth by using

```
> wsw <- llbt.worth(msw)
```

▷ and plot results by

```
> plotworth(wsw)
```



Preferences
estimated worth
0.1
0.2
0.3
0.4
fem W−no
male W−no
fem W−y
male W−y
LO
PA
MI
SG
BA
ST
LO
PA
SG
BA
ST
MI
PA
MI
LO
BA
ST
SG
LO
PA
BA
SG
MI
ST

Example: CEMS exchange programme

- We are interested if universities with a common attribute can be regarded as a group having the same rank

- consider the attribute LAT (with two levels):
- universities PA, MI, BA with latin language: $\text{LAT} = 1$

- universities LO, SG, ST no latin language: $\text{LAT} = 0$

The values for LAT are given as follows:

Objects	LO	PA	MI	SG	BA	ST
LAT	0	1	1	0	1	0

Function: `llbt.design()`

▷ generating the design matrix into data frame

```
> des <- llbt.design(cpc, 6, objnames = cities, cov.sel = c("SEX",
+     "WORK"))
```

▷ categorical subject covariates must be declared as `factor()`

```
> des$SEX <- factor(des$SEX)
> des$WORK <- factor(des$WORK)
```

▷ declare **object covariate:**
reparameterizing the objects (cf. LLTM)

```
> LAT <- c(0, 1, 1, 0, 1, 0)
> objects <- as.matrix(des[6:11])
> mLAT <- objects %*% LAT
```

▷ fit model using standard R function gnm()

gnm() generalised nonlinear models (Turner, Firth)

▷ fit a specific model:
e.g. different preference scales for SEX
but Latin cities (mLAT) combined with WORK

```
> mdsLw <- gnm(y ~ LO + PA + MI + SG + BA + ST + (LO + PA +
+     MI + SG + BA + ST):SEX + mLAT:WORK + g1, elim = mu:SEX:WORK,
+     family = poisson, data = des)
```

Note: g1 is the undecided parameter

Remarks

1. it is assumed that the decisions are independent!
(may be not reasonable)

2. missing values (NA) can occur in the comparisons
just reduce the number of respondents N_{ij}
but no missing values are allowed in the subject covariates

3. the number of rows of the design matrix is:

number of comparisons $\times$
number of possible decisions (response categories) $\times$
number of subject groups

Paired Comparison Pattern Models

- different approach
  but includes all extensions mentioned so far
- more general concerning further extensions
  ▷ pattern models maintain information of
  all individual responses to PC
  ▷ as opposed to LLBT-models, which are marginal models

▷ we model the complete responses $\mathbf{Y}$ simultanously

$$\mathbf{Y} = (Y_{12}, Y_{13}, \ldots Y_{J-1,J})$$

What are paired comparison response patterns?

comparison	(12)	(13)	(23)	...
response	$(1 \succ 2)$	$(3 \succ 1)$	$(2 \succ 3)$	...
random variable	Y_{12}	Y_{13}	Y_{23}	...

The BT Model as a Pattern Model

$$Y_{jk} = \begin{cases} 1 & \text{if object } O_j \text{ is preferred to } O_k \quad (j \succ k) \\ -1 & \text{if object } O_k \text{ is preferred to } O_j \quad (k \succ j) \end{cases}$$

$$P(j \succ k) = P(Y_{jk} = 1) = c \left(\frac{\sqrt{\pi_j}}{\sqrt{\pi_k}} \right)^{y_{jk}}$$

the probability for a specific response pattern e.g. (1, 1, 1) which means $(1 \succ 2)$, $(1 \succ 3)$, $(2 \succ 3)$ is given by:

$$p(1,\,1,\,1\,) = \delta \left(\frac{\sqrt{\pi_1}}{\sqrt{\pi_2}} \right) \left(\frac{\sqrt{\pi_1}}{\sqrt{\pi_3}} \right) \left(\frac{\sqrt{\pi_2}}{\sqrt{\pi_3}} \right)$$

the log-linear pattern model can be written as:

$$\boxed{\ln m(1,\,1,\,1) = \ln \delta + 2\lambda_1 - 2\lambda_3}$$

- all possible patterns are `number of responses (2)`$^{\binom{J}{2}}$ (if no undecided)

Dependencies

one important feature of the pattern models is

- we can give up the (unrealistic) assumption of independent decisions
- we assume that dependencies between responses come from repeated evaluation of the same objects in PC
  comparing (j with k) and comparing (j with l)
  the assessment of common object j might be similar in both comparisons

we can now include dependence terms of the form:

$$\boxed{\theta_{(jk),(jl)}}$$

for pairs of comparisons with one object in common

What makes a good teacher ?

239 education students at Vienna were asked to compare qualities of a good teacher in 2006 through a complete paired comparison experiment
Quality of the teachers are:

ST Structure of instruction
CM Class Management: productive environment - not wasting time
AC Activity: Success in getting students to participate
SU Support: Looking after every single pupil

▷ subject covariates
SEX gender (1 = female) (2 = male)
SCH school (1 = secondary) (2 = vocational) (3 = university)

- no undecided
- but missing values (`NA`)

fit basic model using `pattPC.fit()`

```
> mtp <- pattPC.fit(teacher4, nitems = 4, undec = F, ia = T,
+     formel = ~1, elim = ~SEX * SCH, obj.names = it4)
```

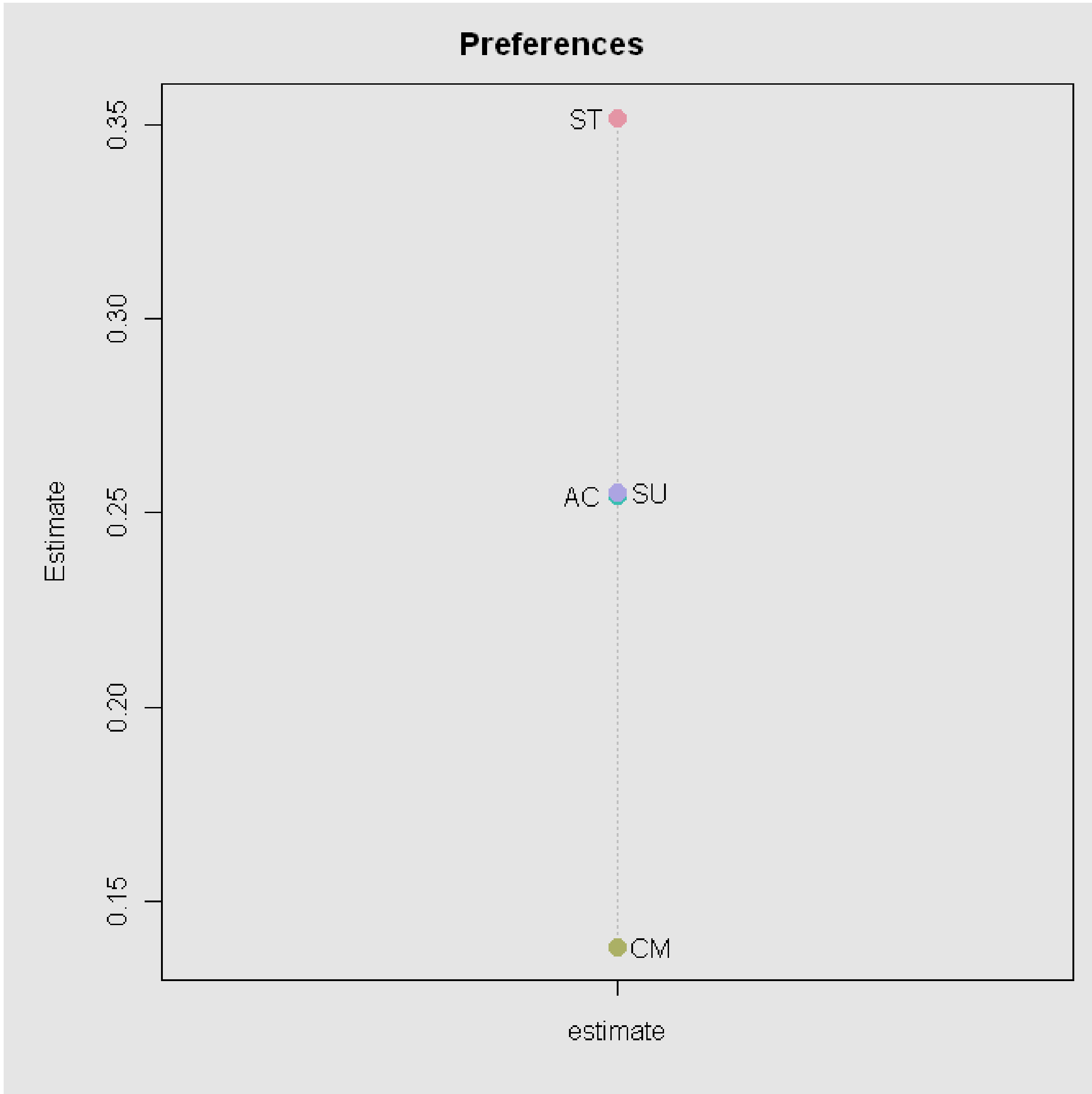
Options = {

teacher4	data.frame
nitems = 4	4 items
undec = F	no undecided
ia = T	all possible dependencies
formel =∼ 1	model is ST+CM+AC+SU
elim =∼ SEX ∗ SCH	defines maximal table
obj.names = it4	names of items

some other Options: ▷ see ?pattPC.fit

Calculate worth and plot using `patt.worth()`, `plotworth()`

```
> wp <- patt.worth(mtp)
> plotworth(wp)
```



Preferences
ST
AC
SU
CM
0.35
0.30
0.25
0.20
0.15
Estimate
estimate

Treatment of Missing Values in Pattern Models

- each different missing pattern gives a different design matrix (smaller than design matrix for non-missing data)
- likelihood is computed for each of these "different" tables – "individual" contributions
- total likelihood (which is then maximised) is the sum of all the "individual" contributions

implemented in `prefmod`

- in `pattPC.fit()` (and in all `patt*.fit()` functions)
- computationally demanding escpecially with large tables and many different missing value patterns
- rough check for "not ignorable" missing use option: `NItest = T`

Example: Rankings

Vargo (1989) collected a ranking data set
which was analysed by Critchlow, Fligner (Psychometrika, 1991)

- 32 judges were asked to rank
  four salad dressings according tartness.

- A low rank means very tart.

salads A - D have varying concentrations
the four pairs of concentrations of acetic and gluconic acid are:
`A = (.5, 0), B = (.5, 10.0), C = (1.0, 0),` **and** `D = (0, 10.0)`

Response Format: Rankings

full rankings:

- people are asked to rank objects (items) regarding a certain aspect (tartness)
- all possible pairs are constructed afterwards
- no undecided category !

ordinal response formats are transformed into paired comparisons

- resulting PCs are called derived PC patterns
- no intransitive patterns possible
- no dependencies

Transformation: Ranking to PC

Data				comparison		
R	G	B	Response	RG	RB	GB
1	2	3	R>G>B	1	1	1
1	3	2	R>B>G	1	1	-1
-	-	-	-	1	-1	1
2	3	1	B>R>G	1	-1	-1
2	1	3	G>R>B	-1	1	1
-	-	-	-	-1	1	-1
3	1	2	G>B>R	-1	-1	1
3	2	1	B>G>R	-1	-1	-1

- number of possible patterns is $3! = 6$ compared to $2^{\binom{3}{2}} = 8$

Pattern Model: Rankings

The probability for the ranking $R = 2$, $G = 3$, $B = 1$ transformed into pattern 1, -1, -1 is given by:

$$p(s_k) \Rightarrow p(y_{12}, y_{13}, y_{23}) = \delta \left(\frac{\sqrt{\pi_1}}{\sqrt{\pi_2}}\right)^1 \left(\frac{\sqrt{\pi_1}}{\sqrt{\pi_3}}\right)^{-1} \left(\frac{\sqrt{\pi_2}}{\sqrt{\pi_3}}\right)^{-1}$$

$$p(2,3,1) \Rightarrow p(s_4) = p(1,-1,-1) = \delta \left(\frac{\sqrt{\pi_1}}{\sqrt{\pi_2}}\right) \left(\frac{\sqrt{\pi_3}}{\sqrt{\pi_1}}\right) \left(\frac{\sqrt{\pi_3}}{\sqrt{\pi_2}}\right)$$

The log expected number for the ranking can be rewritten as

$$\boxed{\ln m(1,-1,-1) = \ln \delta - 2\lambda_2 + 2\lambda_3}$$

direct estimation: **Function:** `pattR.fit()`

▷ fit basic pattern model for rankings

```
> salmod <- pattR.fit(salad, nitems = 4)
> summary(salmod)
```

Calculate worth and plot using `patt.worth()`, `plotworth()`

OR

▷ fit model using design matrix approach
`patt.design()` and use `glm()` or `gnm()`

```
> saldes <- patt.design(salad, nitems = 4, resptype = "ranking")
> salmod2 <- glm(y ~ A + B + C + D, family = poisson, data = saldes)
```

▷ to fit object covariates use design matrix approach

Example: Ratings

we used a data set collected by the British Household Panel Study in 1996 where we have chosen three Likert items which ask respondents for their concern about:

- the destruction of the ozone layer (OZ)
- the high rate of unemployment (UN)
- declining moral standards (MO)

the possible answers are:

- A great deal 1
- A fair amount 2
- Not very much . . . 3
- Not at all 4

low numbers mean a high concern and higher number lower concern!

Transformation: Ratings to PC

for example the Likert response pattern was

$OZ = 1$, $UN = 4$, $MO = 4$

we have 3 items and therefore 3 comparisons:
(12) =(OZ, UN) (13) =(OZ, MO) (23) =(UN, MO)

- as OZ $\succ$ UN we assign $y_{12} = 1$
- as OZ $\succ$ MO we assign $y_{13} = 1$
- as UN $=$ MO we assign $y_{23} = 0$ which is undecided

so we get the following (derived) PC pattern: 1, 1, 0

Pattern Model: Ratings

the probability for the rating $OZ = 1$, $UN = 4$, $MO = 4$ transformed into pattern (1, 1, 0) is given by:

$$p(1,1,0) = \delta \left(\frac{\sqrt{\pi_1}}{\sqrt{\pi_2}}\right) \left(\frac{\sqrt{\pi_1}}{\sqrt{\pi_3}}\right) u_{23}$$

the log expected number for the rating can be rewritten as

$$\ln m(1,1,0) = \ln \delta + 2\,\lambda_1 - 1\lambda_2 - 1\lambda_3 + \gamma_{23}$$

where γ is the undecided parameter

Transformation: Rating to PC

restricted example for 3 items, only 2 response categories

e.g., concern yes= 1 and concern no= 2

Rating patterns			derived PC-patterns			unique PC-patterns		
i_1	i_2	i_3				y_{12}	y_{13}	y_{23}
1	1	1	0	0	0	0	0	0
1	1	2	0	1	1	0	1	1
1	2	1	1	0	-1	1	0	-1
1	2	2	1	1	0	1	1	0
2	1	1	-1	-1	0	-1	-1	0
2	1	2	-1	0	1	-1	0	1
2	2	1	0	-1	-1	0	-1	-1
2	2	2	0	0	0			

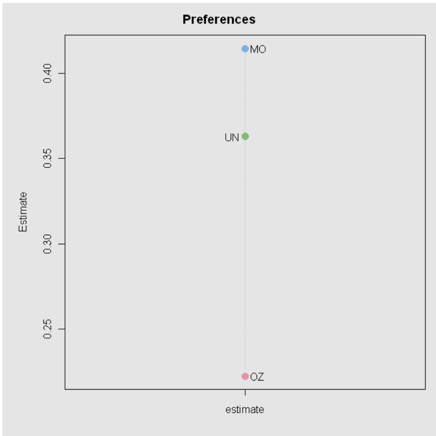
▷ for 3 items only 7 possible patterns (instead of $9 = 3^3$ possible patterns) direct estimation: **Function:** `pattL.fit()`

▷ fit basic pattern model for ratings

```
> t3dat <- read.table("t3dat.dat", header = TRUE)
> lm1 <- pattL.fit(t3dat, 3, undec = T, elim = ~sex * age4k)
```

Calculate worth and plot using `patt.worth(), plotworth()`

```
> w1 <- patt.worth(lm1)
> plotworth(w1)
```



Overview of main prefmod functions

Response	Model	Data	Designmatrix	Estimation	Notes
Real PCs	LLBT	Data	`llbt.design()`	`glm(), gnm()`	1,2,(3),4,5
		Data	`llbt.design()`	`llbt.fit()`	1,3,4,5
		Data	$\longrightarrow$	`llbtPC.fit()`	1,(3),(5),7
	Pattern	Data	`patt.design()`	`glm(), gnm()`	2,4,(5),6
		Data	$\longrightarrow$	`pattPC.fit()`	1,(5),6,7
Ranking	Pattern	Data	`patt.design()`	`glm(), gnm()`	2,4,(5)
		Data	$\longrightarrow$	`pattR.fit()`	1,(3),5,7
Rating	Pattern	Data	`patt.design()`	`glm(), gnm()`	2,4,(5),6
		Data	$\longrightarrow$	`pattL.fit()`	1,5,6,7

(1) NAs
(2) R standard output
(3) larger number of objects
(4) object-specific covariates
(5) continuous subject covariates
(6) dependencies
(7) worth matrix, worth plot

Further Extensions in prefmod()

- multidimensional PC pattern models
  when objects are compared on more than one attribute
- repeated evaluation of the same objects by the same judges (panel data)
- mixture models (latent class) for all extensions
- further response formats
  e.g. partial rankings, *piling*, *best to worst scaling*
- combination of the various options mentioned

LLBT - Models
Dittrich, R., Hatzinger, R., and Katzenbeisser, W. (1998).
Journal of the Royal Statistical Society, Series C.

PC-Pattern Models - Dependencies
Dittrich, R., Hatzinger, R., and Katzenbeisser, W. (2002).
Computational Statistics and Data Analysis.

Multidimensional PC-Pattern Models
Dittrich, R., Francis, B., Hatzinger, R., and Katzenbeisser, W. (2006).
Mathematical Social Sciences.

(Likert) Rating Pattern Models
Dittrich, R., Francis, B., Hatzinger, R., and Katzenbeisser, W. (2007).
Statistical Modelling.

Temporal Dependence in Longitudinal Paired Comparisons
Dittrich, R., Francis, B., Hatzinger, R. and Katzenbeisser, W. (2008).
Research Report Series.

Ranking Pattern Models - latent classes
Francis, B., Dittrich, R., and Hatzinger, R. (2009).
under revision.

Missing Values in Pattern Models
Dittrich, R., Francis, B., Hatzinger, R., and Katzenbeisser, W. (2009).
under revision.

Partial-Ranking Pattern Models
Darbic, M., Hatzinger, R. (2009).
In: Präferenzanalyse mit R. eds: Hatzinger, Dittrich, Salzberger